

Determination of Substitutional Sites in Heterocycles from the Topological Analysis of the Electron Localization Function (*ELF*)

FRANCK FUSTER, ALAIN SEVIN, BERNARD SILVI

Laboratoire de Chimie Théorique (UMR-CNRS 7616), Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris cédex 05, France

Received 9 August 1999; accepted 7 December 1999

ABSTRACT: The topological analysis of the electron localization function *ELF* has been carried out on five-membered (C_4H_4NH , C_4H_4PH , C_4H_4O , C_4H_4S) and six-membered (C_5H_5N , C_5H_5P) heterocycles. The bonding in these molecules is discussed on the basis of the valence basin populations. It is shown that the values of the *ELF* function at the $(3, -1)$ critical points between disynaptic basins related to a given center provide a criterion to determine substitutional sites.

© 2000 John Wiley & Sons, Inc. J Comput Chem 21: 509–514, 2000

Keywords: heterocycles; topological analysis; electron localization function (*ELF*)

Introduction

Since its introduction in 1990 by Becke and Edgecombe,¹ the electron localization function *ELF* has been acknowledged as a very useful tool for the analysis of wave functions.² In 1994, Silvi and Savin³ generalized to *ELF* the topological approach pioneered by Richard Bader.⁴ The topological analysis of *ELF* has then been applied to a large number of systems for which it provides a very “chemical” picture of the bonding,^{5–16} as well as of reactive processes.^{17–20} The *ELF* function is a local indication of the Pauli repulsion strength²¹ and, therefore,

its gradient field enables to partition the molecular space into basins of attractors that correspond to VSEPR electron pair domains.^{22, 23} Indeed, the *ELF* analysis offers a unified description of the bonding and of the reactivity, which is very close to the ideas of electron pairing. In the development of the *ELF* analysis we have been led to introduce several new concepts and a specific vocabulary. All the necessary definitions can be found in refs. 3, 5, 9, and 17.

In recent applications on protonation sites in bases²⁴ and on the aromatic electrophilic substitution²⁵ we have shown that a convenient use of the information carried by *ELF* can be used to determine the reactive centers of a molecule. The aim of this contribution is to show that the techniques

Correspondence to: B. Silvi; e-mail: silvi@lct.jussieu.fr

used to study the protonation and the aromatic substitution can be generalized to some aspects of the heterocycle chemistry. The investigated systems are, on the one hand, four five-membered molecules C_4H_4NH , C_4H_4PH , C_4H_4O , and C_4H_4S , and on the other hand, two six-membered cycles C_5H_5N and C_5H_5P .

Results and Discussion

The calculations have been performed with the standard 6-31G** basis set^{26–28} at the hybrid Hartree–Fock density functional Becke3LYP^{29–32} levels using the Gaussian 94 package.³³ The Becke-3LYP method used with the 6-31G** basis set yields satisfactory optimized geometries for the investigated compounds. Moreover, the topology of the *ELF* function being very stable with respect to the nature of the basis set, the improvement of this latter by additional diffuse or polarization functions has no effect on the number of basins, on their nature, and a very little one on the basin properties. The Laplacian of the charge density has been calculated with the AIMPAC series of programs.³⁴ The *ELF* analysis has been performed with the ToPMoD programs developed in our laboratory,^{35,36} whereas the graphical representations have been done with the SciAn software.³⁷

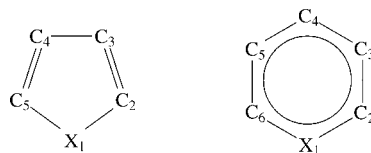
The *ELF* picture of the orientational effects in the electrophilic aromatic substitution essentially relies upon the analysis of the bifurcations of the localization domains. An *f*-localization domain is a volume bounded by a closed isosurface of the *ELF* function $\eta(\mathbf{r}) = f$. It is said to be reducible if it contains more than one attractor, irreducible otherwise.³ At a low enough value of the bounding isosurface $\eta(\mathbf{r})$, one gets a unique localization domain that contains all the attractors. When $\eta(\mathbf{r})$ is increased, this single domain is split into new localization domains. This process, called reduction of localization,⁵ is continued until all domains are irreducible. The reduction of reducible domains is a criterion of discrimination between basins. The reduction of a reducible localization domain occurs at a critical value of the bounding isosurface, over which the domain is split into domains containing fewer attractors. The localization domains are then ordered with respect to the *ELF* critical values yielding bifurcations. These values are precisely those of the localization function at the $(3, -1)$ critical points lying on the separatrices between valence basins. The hierarchy of the bifurcation can be visualized by a tree diagram. The highest bifurcation occurs between the valence

basins related to the most electronegative atom of a molecule, and therefore, the bifurcation diagrams give an indication of the relative electronegativities and, therefore, of the reactivity of the atoms in a molecule. For example, in the case of aniline,²⁵ the highest bifurcation within the phenyl group occurs for the carbon in the *ortho* position with respect to the substituent it corresponds to the last reduction of the “aromatic domain,” which initially contains all the attractors of the $V(C_i, C_{i+1})$ basins. The successive reductions of the aromatic domain take place at the $(3, -1)$ critical points between the $V(C_{i-1})$ and $V(C_i, C_{i+1})$ basins at the level of the carbons labeled by C_i and for the value referred to as $\eta(C_i; NH_2)$ ($\eta(C_i; S)$ in the general case of a monosubstituted benzene C_6H_5S). The electrophilic substitution positional indices of the carbon in position *c* with respect to the substituent:

$$RI_c(S) = \eta(C_i; S) - \eta(C_i; H) \quad (1)$$

are differences of the *ELF* values at equivalent topological invariants in the S-derivative and in benzene. These indices are additive and, therefore, the regioselectivity in a polysubstituted compound can be inferred from the indices of monosubstituted analogs.

Figure 1 displays the localization domains of the C_4H_4X ($X = NH, PH, O, S$), C_5H_5N , and C_5H_5P molecules. The value of the defining bounding isosurface has been chosen for each molecule to clearly indicate the place of the highest bifurcation, which divides the bonding valence domains. For all systems the core of the heteroatom is surrounded by a single localization domain, which contains all the attractors of the valence basins linked to this center. The heteroatoms tend to keep their valence shell structure at larger *ELF* values than the carbons. The usual convention³⁸ has been adopted to label the carbons:



In pyrrole, furane, and thiophene the highest bifurcations around a carbon occur at two symmetry related $(3, -1)$ critical points located on the separatrices of the $V(C_2, C_3)$ and $V(C_2, H)$ basins on the one hand, and on that of $V(C_4, C_5)$ and $V(C_5, H)$ on the other hand. The reduction of localization of pyrrole (Fig. 2) is representative of this series of molecules. In the case of C_4H_4PH , the protonated disynaptic domains $V(C_2, H)$ and $V(C_5, H)$ are detached

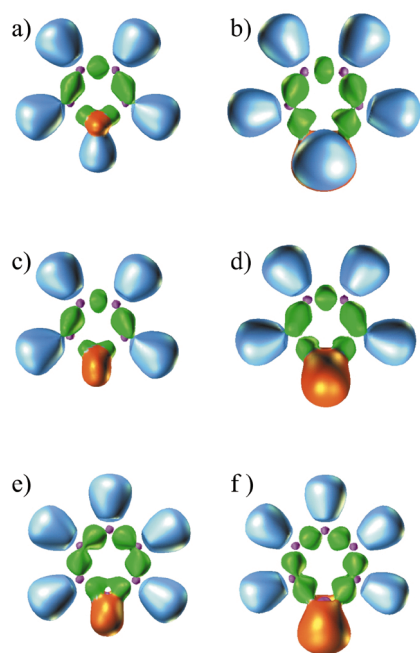


FIGURE 1. Localization domains of the investigated heterocycles. (a) C_4H_4NH , (b) C_4H_4PH , (c) C_4H_4O , (d) C_4H_4S , (e) C_5H_5N , (f) C_5H_5P . This figure presented here in black and white is available in color on the World Wide Web at <http://journals.wiley.com/jcc/>. Color code: magenta = core, red = monosynaptic, blue = protonated disynaptic, green = disynaptic.

of respectively $V(C_2, C_3)$ and $V(C_4, C_5)$ before the $V(C_2, P)$ and $V(C_5, P)$ ones.

The topology of the six-membered ring is quite similar to that of monosubstituted benzenes. In pyridine, like in nitrobenzene, the ring is first opened at the level of the *para* and *ortho* carbon and after at C_3 – C_5 symmetrically. C_5H_5P behaves more like aniline, because the ring highest bifurcation is in *ortho* rather than in *meta*.

The valence basin populations reported in Tables I and II shed light on the bonding in these molecules. In the standard resonance picture of the five-membered rings the lone pair(s) of the heteroatom is(are) assumed to participate to the delocalization. Accordingly, one can expect that the population of the corresponding $V(X)$ basins should be less than $2 e^-$. This happens to be verified by the ELF analysis only in the case of pyrrole. Pyrrole is a planar molecule in which the nitrogen lone pair gives rise to two symmetrical monosynaptic basins $V_1(N)$ and $V_2(N)$, each of them with a population of $0.7 e^-$ (gathered in Table I). The low population of the lone-pair basins is a consequence of the planar geometry of the tricoordinated nitrogen rather than

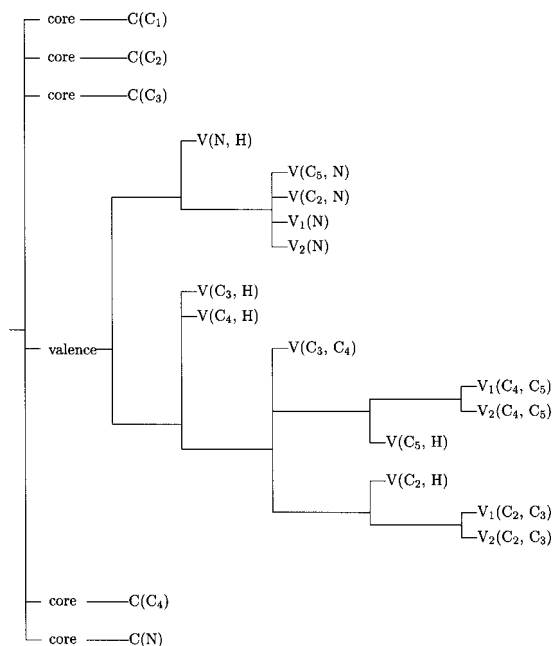


FIGURE 2. Localization domain reduction tree diagram of pyrrole.

of delocalization. In C_4H_4PH , the C–P distances are large enough to enable the buckling of the ring, and the phosphorus lone pair is represented by a single monosynaptic basin $V(P)$ with a population of $2.12 e^-$. Moreover, as pointed out by Kutzelnigg,³⁹ pyramidalization around second row elements is

TABLE I. Valence Basin Population, ELF Values at Bifurcation Critical Points and Charge Density Laplacian at Secondary Charge Concentrations of Five-Membered Heterocycles C_4H_4X .

	C_4H_4NH	C_4H_4PH	C_4H_4O	C_4H_4S
$V(C_2, H)$	2.17	2.12	2.18	2.15
$V(C_3, H)$	2.12	2.12	2.12	2.11
$V(C_2, X)$	2.02	2.20	1.57	1.86
$V(C_2, C_3)$	3.46	3.37	3.58	3.41
$V(C_3, C_4)$	2.48	2.12	2.33	2.43
$V(X)$	1.40	2.12	4.21	4.04
$V(X, H)$	2.06	1.92		
$\eta(C_2)$	0.7134	0.7029	0.7294	0.6960
$\eta(C_3)$	0.7047	0.6586	0.6895	0.6777
$\nabla^2 \rho(C_2)$	−0.239	−0.132	−0.136	−0.136
$\nabla^2 \rho(C_3)$	0.081	0.088	0.084	0.084

The symmetrical $V(X)$ and $V(C_2, C_3)$ have been merged in single superbasins.

TABLE II.
Valence Basin Population, ELF Values at Bifurcation
Critical Points and Charge Density Laplacian at
Secondary Charge Concentrations of Six-Membered
Heterocycles C₅H₅X.

	C ₅ H ₅ N	C ₅ H ₅ P
V(C ₂ , H)	2.17	2.10
V(C ₃ , H)	2.11	2.11
V(C ₄ , H)	2.11	2.11
V(C ₂ , X)	2.31	2.50
V(C ₂ , C ₃)	2.95	2.75
V(C ₃ , C ₄)	2.77	2.76
V(X)	2.69	2.88
$\eta(\text{C}_2)$	0.6546	0.6856
$\eta(\text{C}_3)$	0.6712	0.6573
$\eta(\text{C}_4)$	0.6544	0.6647
$\nabla^2\rho(\text{C}_2)$	0.091	-0.214
$\nabla^2\rho(\text{C}_3)$	-0.172	0.095
$\nabla^2\rho(\text{C}_4)$	0.096	0.092

made easier than for the corresponding first row ones by the relative sizes of the *s* and *p* orbitals.

In thiophene and furan, the planar symmetry is compatible with the presence of two V(X) basins, each of these latter having a population of respectively 2.02 and 2.10 e^- . Nevertheless, the analysis of the contributions of the other basins to the variance of the V(X) basin population⁹ provides some indications on the delocalization,⁴⁰ which mostly concerns the adjacent valence basins, namely V(C₂, X), V(C₅, X) and V(X, H), or the second V(X) according to the molecule. In all molecules, the C₂—C₃ and C₄—C₅ bonds have a clear double-bond character testified by the presence of two symmetrical disynaptic basins and by total populations larger than 3 e^- . The population of the V(C₃, C₄) basin is always larger than 2 e^- due to a charge transfer from V(C₂—C₃) and V(C₄—C₅), whereas the V(C₂, X) and V(C₅, X) have population either greater than 2 e^- (pyrrole and C₄H₄PH) or less than 2 e^- (furan and thiophene). The V(C₃, C₄) basin is always involved in the delocalization with the adjacent disynaptic basins. With respect to the nature of the heteroatom, the V(C₃, C₄) population increases with its electronegativity, whereas that of the V(C, X) follows the opposed trend within a given column: V(C, P) > V(C, N) and V(C, S) > V(C, O) but V(C, S) < V(C, N) because of a weaker delocalization of the lone pairs towards V(C, X) in furan and thiophene.

The basin populations in pyridine and phosphidine indicate that these molecules have a more pronounced aromatic character than the five-

membered rings. On the one hand, all the attractors of the disynaptic basins between the ring atoms are located in the molecular plane; on the other hand, the populations of these basins range between 2.31 and 2.95 e^- . The larger delocalization in these molecules takes place essentially between adjacent disynaptic basins for which the variance contributions are about twice that calculated in five-membered rings. The rather large charge transfer towards the heteroatom lone pair basin (σ -donation) makes this ring electron deficient with respect to benzene.

Tables I and II also report the ELF value at the (3, -1) critical point corresponding to the highest bifurcation around each carbon and the value of $\nabla^2\rho$ at the out-of-plane saddle points corresponding to the secondary charge concentrations.⁴¹ In all the five-membered rings the highest $\eta(r_c)$ as well as the negative $\nabla^2\rho$ are found for the carbons C₂ and C₅ (position α), which are the favored sites of the electrophilic substitution. This result is consistent with the predictions of other theoretical methods such as frontier orbitals⁴² and with the experimental results.³⁸ In the case of pyridine, one has to consider the nucleophilic rather than the electrophilic substitution, this latter being hampered by the electron deficiency of the ring. The question of transferability of our approach to the case of nucleophilic aromatic substitution can be risen. Obviously, in this study we have mostly focussed our attention on hydrogen-substituted ring carbons. When dealing with nucleophilic attack, the emphasis would have to be focussed on the ELF value at carbons bearing the potential leaving group. But the same type of calculation would be used as well without any significant methodological change. The critical values for the C₂ and C₄ carbon are almost equal or less than that of C₃. Experimentally, in the Chichibabin reaction the nucleophilic attack occurs on C₂ or C₆ unless both positions are occupied. In such cases substitution can occur on C₄.³⁸ It is worth noting that the frontier electron population of C₂ is larger than of C₄ by 0.15,⁴² indicating this latter as the nucleophilic attack site. Assuming equal probabilities for C₂, C₄, and C₆, it is finally the *ortho* position that is favored because of its multiplicity. In phosphidine, our calculation predicts that the nucleophilic attack might occur on the *meta* position.

For substituted heterocycles it is possible to define positional substitution indices according to eq. (1). In the case of the aromatic substitution, these indices are additive, enabling a straightforward estimate for polysubstituted molecules from monosubstituted values. Table III reports the values calculated for the 2-fluoropyrrole, 3-fluoropyrrole,

TABLE III.
Substitution Positional Indices of *ortho*- and *meta*-Fluoropyrroles and of the Difluoropyrrole Isomers.

	C ₂	C ₃	C ₄	C ₅
2-C ₄ H ₃ FNH		0.020	−0.010	0.013
3-C ₄ H ₃ FNH	0.021		0.011	−0.001
4-C ₄ H ₃ FNH	−0.001	0.011		0.021
5-C ₄ H ₃ FNH	0.013	−0.010	0.020	
2,3-C ₄ H ₂ FNH		0.006 (0.001)	0.012 (0.012)	
2,4-C ₄ H ₂ FNH		0.031 (0.031)		0.033 (0.034)
2,5-C ₄ H ₂ FNH		0.012 (0.010)	0.012 (0.010)	
3,4-C ₄ H ₂ FNH	0.019 (0.020)			0.019 (0.020)

() = estimated values.

and for the various difluoropyrroles. Except for the C₃ index of the 2,3-difluoropyrrole for which there is a small discrepancy between the calculated and estimated values, the additivity of the substitution positional indices is quite perfectly verified.

Conclusion

The topological analysis of the electron localization function provides a description of the bonding in heterocycles that enables location of the centers of electrophilic and nucleophilic attacks. In the five-membered rings, except pyrrole, the heteroatom lone pair(s) appear(s) to have populations slightly larger than $2e^-$, which is not consistent with the standard resonance picture unless a strong σ back-donation compensates the π -donation. In furan, thiophene, and phosphole, this back-donation is possible because the lone-pair axis is not perpendicular to the molecular plane. In pyrrole, the back-donation is not possible, and the lone pair is depleted. The value of the function at the (3, −1) critical points, which correspond to the highest reduction of the localization domains, appears to be a suitable token of positional reactivity.

Our determination of electrophilic (and conversely nucleophilic) attack sites is clearly semi-quantitative, but it nevertheless remains of chemical significance. A short comparison with frontier MO and Atoms in Molecules theories can be made. On the one hand, a well-driven perturbation analysis may yield the relative sizes of the coefficients in the HOMO of a given monosubstituted π framework. However, it has often been shown that several MOs have to be considered in many cases (underlying MO control). Moreover, it becomes quite delicate to treat by perturbation the case of a polysubstituted

ring. These severe limitations are easily overcome at a rather low cost by our method through the definition of the positional substitution indices that are additive. On the other hand, the values of $\nabla^2\rho$ at the secondary charge concentrations follow the same trend as the ELF values at the carbon highest bifurcation points.

Acknowledgments

We are grateful to professor R. F. W. Bader for a copy of the AIMPAC package.

References

1. Becke, A. D.; Edgecombe, K. E. *J Chem Phys* 1990, 92, 5397.
2. Savin, A.; Nesper, R.; Wengert, S.; Fässler, T. F. *Angew Chem Int Ed Engl* 1997, 36, 1809.
3. Silvi, B.; Savin, A. *Nature* 1994, 371, 683.
4. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford Univ. Press: Oxford, 1990.
5. Savin, A.; Silvi, B.; Colonna, F. *Can J Chem* 1996, 74, 1088.
6. Llusar, R.; Beltrán, A.; Andrés, J.; Silvi, B. *J Phys Chem* 1995, 99, 12483.
7. Alikhani, M. E.; Bouteiller, Y.; Silvi, B. *J Phys Chem* 1996, 100, 16092.
8. Silvi, B.; Savin, A.; Wagner, F. R. In *Modelling of Minerals and Silicated Materials*; Silvi, B.; Arco, P. D., Eds.; Kluwer Academic Publishers: Dordrecht, 1977, p. 179, Vol. 15.
9. Noury, S.; Colonna, F.; Savin, A.; Silvi, B. *J Mol Struct* 1998, 450, 59.
10. Alikhani, M. E.; Silvi, B. *J Comput Chem* 1999, 19, 1205.
11. Beltrán, A.; Andrés, J.; Noury, S.; Silvi, B. *J Phys Chem A* 1999, 103, 3078.
12. Fourré, I.; Silvi, B.; Chaquin, P.; Sevin, A. *J Comput Chem* 1999, 20, 897.
13. Llusar, R.; Beltrán, A.; Andrés, J.; Noury, S.; Silvi, B. *J Comput Chem* 1999, 20, 1517.

14. Fuster, F.; Silvi, B. *Theor Chem Acc* 1999, to appear.
15. Joubert, L.; Picard, G.; Silvi, B.; Fuster, F. *J Electrochem Soc* 1999, 146, 2180.
16. Joubert, L.; Picard, G.; Silvi, B.; Fuster, F. *J Mol Struct (Theochem)* 1999, 463, 75.
17. Krokidis, X.; Noury, S.; Silvi, B. *J Phys Chem A* 1997, 101, 7277.
18. Krokidis, X.; Goncalves, V.; Savin, A.; Silvi, B. *J Phys Chem A* 1998, 102, 5065.
19. Krokidis, X.; Silvi, B.; Alikhani, M. E. *Chem Phys Lett* 1998, 292, 35.
20. Krokidis, X.; Vuilleumier, R.; Borgis, D.; Silvi, B. *Mol Phys* 1999, 96, 265.
21. Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; von Schnering, H. G. *Angew Chem Int Ed Engl* 1991, 30, 409.
22. Gillespie, R. J.; Robinson, E. A. *Angew Chem Int Ed Engl* 1996, 35, 495.
23. Bader, R. F. W.; Gillespie, R. J.; MacDougall, P. J. *J Am Chem Soc* 1988, 110, 7329.
24. Fuster, F.; Silvi, B. *Chem Phys* 1999, to appear.
25. Fuster, F.; Sevin, A.; Silvi, B. *J Phys Chem A* 1999, to appear.
26. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J Chem Phys* 1972, 56, 2257.
27. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; von Ragué Schleyer, P. *J Comput Chem* 1983, 4, 294.
28. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J Chem Phys* 1984, 80, 3265.
29. Becke, A. D. *J Chem Phys* 1993, 98, 5648.
30. Becke, A. D. *Phys Rev* 1988, A38, 3098.
31. Lee, C.; Yang, Y.; Parr, R. G. *Phys Rev* 1988, B37, 785.
32. Miechlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem Phys Lett* 1989, 157, 200.
33. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision D.4*; Gaussian Inc.: Pittsburgh, PA, 1995.
34. Biegler-König, F. W.; Bader, R. F. W.; Tang, T.-H. *J Comput Chem* 1982, 3, 317.
35. Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Topmod package*, 1997.
36. Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Comput Chem* 1999, 23, 597.
37. Pepke, E.; Murray, J.; Lyons, J.; Hwu, T.-Z. *Scian; Supercomputer Computations Research Institute, Florida State University: Tallahassee, FL*, 1993.
38. Streitwieser, A., Jr.; Heathcock, C. H. *Introduction to Organic Chemistry*; McMillan: New York, 1976.
39. Kutzelnigg, W. *Angew Chem Int Ed Engl* 1984, 23, 272.
40. Fradera, X.; Austen, M. A.; Bader, R. F. W. *J Phys Chem A* 1998, 103, 304.
41. Bader, R. F. W.; Chang, C. *J Phys Chem* 1989, 93, 2946.
42. Fleming, I. *Frontier Orbitals and Organic Chemical Reactions*; John Wiley & Sons Inc.: New York, 1976.